Six Generalized Schur Complements

C. A. Butler
Daniel H. Wagner Associates
Hampton Office
Suite 301, 27 West Queens Way
Hampton, VA 23669

and

T. D. Morley
School of Mathematics
Georgia Institute of Technology
Atlanta, Georgia 30332
BITNET: MA201TM AT GITVM1

Submitted by David H. Carlson

ABSTRACT

We give a unified treatment of equivalence between some old and new generalizations of the Schur complement of matrices.

1. INTRODUCTION

Let X be a finite dimensional real or complex vector space with inner product $(\cdot,\cdot)$. Let $A: X \to X$ be a linear operator. Given a subspace $S \subseteq X$, we may write

$$a = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11}: S \to S$, $A_{12}: S^{\perp} \to S$, $A_{21}: S \to S^{\perp}$, and $A_{22}: S^{\perp} \to S^{\perp}$. Here $S^{\perp}$ denotes the orthogonal complement of the subspace S.

If the operator A_{22} is invertible, then the *Schur complement of* A *to the subspace* S is defined by

$$S(A) = A_{11} - A_{12}A_{22}^{-1}A_{21};$$

52 Vanderbilt Ave., New York, NY 10017 0024-3795/88/$3.50

see [22, 23]. The Schur complement has proved useful in a wide variety of contexts; see for example [7, 10, 18, 19, 24].

In order to maintain compatibility amongst the various definitions below, we extend the Schur complement to X by defining it to be zero on $\mathrm{S}^{\perp}$:

$$\mathrm{S}(A) = \begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & 0 \end{bmatrix}.$$

An operator $A: X \to X$ is termed positive, written $A \geqslant 0$, if $A = A^*$ and $(Ax, x) \geqslant 0$ for all $x \in X$. If A is a linear operator, $\operatorname{range} A$ and $\ker A$ denote its range and null space respectively.

The following, which may be found in [2], is an elementary exercise in linear algebra.

LEMMA 1. *Let $A \geqslant 0$. given a subspace* S, *partition A as before*:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

Then $\operatorname{range} A_{21} \subseteq \operatorname{range} A_{22}$ *and* $\ker A_{22} \subseteq \ker A_{12}$.

Given an operator $A: X \to Y$, the Moore-Penrose pseudoinverse $A^{\dagger}$ is the unique operator from Y to X such that

$$A^{\dagger}y = \begin{cases} \text{the unique } x \in (\ker A)^{\perp} & \\ \quad \text{with } Ax = y & \text{if} \quad y \in \operatorname{range} A, \\ 0 & \text{if} \quad y \in (\operatorname{range} A)^{\perp}. \end{cases}$$

A class of generalized Schur complements not considered here are those based on new partial orders such as the "minus" partial order. For a survey of these generalized Schur complements, see [8]. We also do not consider generalizations for pairs of subspaces; see for example the shorted operator of Mitra et al. [17].

2. SIX GENERALIZED SCHUR COMPLEMENTS

In this section we define six generalized Schur complements. Some of the definitions require hypotheses on the operator A, and some require a bit of proof to show that they are well defined. In all cases $A: X \to X$ is a linear operator, partitioned with respect to a given subspace $S \subseteq X$:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

GENERALIZATION I (Albert [1]; Carlson, Haynsworth, and Markham [9]). Define

$$S_{\mathrm{I}}(A) = \begin{bmatrix} A_{11} - A_{12}A_{22}^{\dagger}A_{21} & 0 \\ 0 & 0 \end{bmatrix}.$$

For the following let $P_{S^\perp}$ denote the orthogonal projection onto $S^\perp$.

GENERALIZATION II (Ando [5]). Suppose that there are matrices M_l and M_r such that

$$P_{S^\perp} M_r = M_r, \qquad M_l P_{S^\perp} = M_l, \tag{1}$$

$$P_{S^\perp} A M_r = P_{S^\perp} A, \quad \text{and} \quad M_l A P_{S^\perp} = A P_{S^\perp}. \tag{2}$$

Then we define

$$S_{\mathrm{II}}(A) = A - AM_r.$$

This definition is independent of the choices of M_l and M_r. We have interchanged S and $S^\perp$ in Ando's original definition to maintain notational consistency with the other definitions.

For the following definition we define a partial order on operators. Let us write $A \geqslant B$ if $A - B$ is positive.

GENERALIZATION III (Krein [15]; Anderson [2]; Anderson and Trapp [4]). Suppose $A \geqslant 0$. Set

$$\mathcal{M}(A, S) = \{ X \geqslant 0 : X \leqslant A, \text{range}\, X \subseteq S \}.$$

Then define

$$S_{\mathrm{III}}(A) = \sup \mathcal{M}(A, S).$$

The proof of the existence of $S_{\mathrm{III}}(A)$ may be found in the above-cited references.

The following definition comes from solving a linear system that arises in electrical circuit theory. See for example Anderson [2] and Anderson, Morley, and Trapp [3]. If $c \in S$, $y \in S^{\perp}$, we write $\begin{bmatrix} c \\ y \end{bmatrix}$ in conformity with the partition of A. If $c \in S$, note that we may write both Ac and $A\begin{bmatrix} c \\ 0 \end{bmatrix}$ for the image of c under A.

GENERALIZATION IV. Suppose $\operatorname{range} A_{21} \subseteq \operatorname{range} A_{22}$ and $\ker A_{22} \subseteq \ker A_{12}$. For each $\begin{bmatrix} c \\ y \end{bmatrix} \in X$ $(c \in S,\ y \in S^{\perp})$ find an $x \in S^{\perp}$ that solves $A_{21}c + A_{22}x = 0$, and define

$$S_{\mathrm{IV}}(A)\begin{bmatrix} c \\ y \end{bmatrix} = A\begin{bmatrix} c \\ x \end{bmatrix}.$$

We will show below that the range and kernel conditions of Generalization IV guarantee that any $x \in S^{\perp}$ that solves $A_{21}c + A_{22}x = 0$ yields the same $A\begin{bmatrix} c \\ x \end{bmatrix}$.

Our penultimate definition is an abstraction of a formula due to Fillmore and Williams [13] and Anderson and Trapp [4] for the shorted operator of a positive operator in infinite dimensions. According to [13] and [4]

$$S(A) = \begin{bmatrix} A_{11} - C^*C & 0 \\ 0 & 0 \end{bmatrix},$$

where C is formally the operator $A_{22}^{1/2\dagger}A_{21}$. We abstract this as

GENERALIZATION V. If there exist operators X, Y, C, and D with

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & CY \\ XD & XY \end{bmatrix}$$

and with $\ker X \subseteq \ker C$ and $\operatorname{range} D \subseteq \operatorname{range} Y$, then we set

$$S_{\mathrm{V}}(A) = \begin{bmatrix} A_{11} - CD & 0 \\ 0 & 0 \end{bmatrix}.$$

This definition is independent of the choices of X, Y, C, D.

For our last definition, we present a geometric version of Generalization IV.

GENERALIZATION VI. Suppose for each $c \in S$ the set

$$\left\{A\begin{bmatrix} c \\ 0 \end{bmatrix} + A(S^{\perp})\right\} \cap S$$

is a singleton. Then define

$$S_{\mathrm{VI}}(A)\begin{bmatrix} c \\ y \end{bmatrix}$$

to be this element.

First some general comments about the above six definitions of generalizations of the Schur complement. Clearly generalization I, since it has no side conditions, exists for any square matrix. However, Generalization I does not satisfy the Haynsworth quotient formula; see [5]. Generalization III, which is known as the *shorted operator*, can be shown (see [4]) to exist for any positive self-adjoint operator on a Hilbert space. A modification of Generalization IV also works in infinite dimensions; see [6].

When we say that two generalizations are equivalent, we mean that the stated side conditions for each generalization imply those for the other generalization, and that the generalized Schur complements determined are equal. The results of this paper can be summarized as:

THEOREM. *Generalizations* II, IV, V, *and* VI *are equivalent. If A is positive, then* $S_{\mathrm{III}}(A) = S_i(A)$ *for* $i = \mathrm{I}, \mathrm{II}, \dots, \mathrm{VI}$. *In any case if* $S_i(A)$ *exists for* $i = \mathrm{II}, \mathrm{III}, \dots,$ *or* VI, *then* $S_i(A) = S_{\mathrm{I}}(A)$.

In what follows, we prove the above in a sequence of propositions.

PROPOSITION 1. *Generalizations* IV *and* VI *are equivalent.*

Proof. Assume that $\operatorname{range} A_{21} \subseteq \operatorname{range} A_{22}$ and $\ker A_{22} \subseteq \ker A_{12}$. Consider the set

$$\left\{A\begin{bmatrix} c \\ 0 \end{bmatrix} + A(S^{\perp})\right\}.$$

If the intersection of this set with S is to be nonempty then we need an $x \in S^{\perp}$ such that

$$A\begin{bmatrix} c \\ x \end{bmatrix} \in S.$$

But

$$A\begin{bmatrix} c \\ x \end{bmatrix} = \begin{bmatrix} A_{11}c + A_{12}x \\ A_{21}c + A_{22}x \end{bmatrix},$$

and $\operatorname{range} A_{21} \subseteq \operatorname{range} A_{22}$ means there exists an $x \in S^{\perp}$ with $A_{22}x = A_{21}(-c)$ or $A_{21}c + A_{22}x = 0$. Moreover, if z also satisfies $A_{21}c + A_{22}z = 0$, then $A_{22}(x-z) = 0$, so $x - z \in \ker A_{22}$, and by our hypothesis $x - z \in \ker A_{12}$. Thus, $A_{12}x = A_{12}z$ and so $A_{11}c + A_{12}x = A_{11}c + A_{12}z$. This shows that

$$\left\{A\begin{bmatrix} c \\ 0 \end{bmatrix} + A(S^{\perp})\right\} \cap S$$

is a singleton and also that

$$S_{\mathrm{IV}}(A)\begin{bmatrix} c \\ y \end{bmatrix} = S_{\mathrm{VI}}(A)\begin{bmatrix} c \\ y \end{bmatrix}.$$

Conversely, the fact that the intersection is nonempty shows that $\operatorname{range} A_{21} \subseteq \operatorname{range} A_{22}$, while the fact that the intersection is a singleton shows that $\ker A_{22} \subseteq \ker A_{12}$. ■

PROPOSITION 2. *If the hypotheses of Generalization* IV *are satisfied, then* $S_{\mathrm{IV}}(A) = S_{\mathrm{I}}(A)$.

Proof. Given $\begin{bmatrix} c \\ y \end{bmatrix}$, let $x = A_{22}^{\dagger}A_{21}(-c)$. Then $A_{21}c + A_{22}x = A_{21}C + A_{22}\ A_{22}^{\dagger}A_{21}(-c)$, but $A_{22}A_{22}^{\dagger}$ is the projection onto $\operatorname{range} A_{22}$, and since $\operatorname{range} A_{21} \subseteq \operatorname{range} A_{22}$ by our hypothesis, we see that $A_{22}^{\dagger}A_{22}A_{21} = A_{21}$.

Thus, $A_{21}c + A_{22}x = 0$, and so

$$S_{\mathrm{IV}}(A)\begin{bmatrix} c \\ y \end{bmatrix} = A\begin{bmatrix} c \\ x \end{bmatrix} = \begin{bmatrix} A_{11}c + A_{12}A_{22}^{\dagger}A_{21}(-c) \\ 0 \end{bmatrix}$$

$$= S_{\mathrm{I}}(A)\begin{bmatrix} c \\ y \end{bmatrix}. \qquad \blacksquare$$

The following proposition may be found (more or less) in [5].

PROPOSITION 3. *Generalizations* IV *and* II *are equivalent.*

Proof. Assume the hypotheses of Generalization IV hold. Recall that this implies $A_{22}A_{22}^{\dagger}A_{21} = A_{21}$. Now define

$$M_r := \begin{bmatrix} 0 & 0 \\ A_{22}^{\dagger}A_{21} & I \end{bmatrix}$$

and

$$M_l := \begin{bmatrix} 0 & A_{12}A_{22}^{\dagger} \\ 0 & I \end{bmatrix}.$$

Then

$$AM_r = \begin{bmatrix} A_{12}A_{22}^{\dagger}A_{21} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

and

$$M_lA = \begin{bmatrix} A_{12}A_{22}^{\dagger}A_{21} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

The (1,2) entry of M_lA is actually $A_{12}A_{22}^{\dagger}A_{22}$, but $A_{22}^{\dagger}A_{22}$ is the projection onto $(\ker A_{22})^{\perp}$, which by our hypothesis contains $(\ker A_{12})^{\perp}$, and so $A_{12}A_{22}^{\dagger}A_{22} = A_{12}$.

From these calculations, one sees that the conditions of Generalization II are satisfied and that

$$S_{\mathrm{II}}(A) = A - AM_r = \begin{bmatrix} A_{11} - A_{12}A_{22}^{\dagger}A_{21} & 0 \\ 0 & 0 \end{bmatrix}.$$

By Proposition 2, we have $S_{\mathrm{II}}(A) = S_{\mathrm{IV}}(A)$. Conversely, suppose that the conditions of Generalization II are satisfied. If we show that the proper range and kernel containments are satisfied, we will be done, by the first half of this proof. Now the conditions (1) of Generalization II imply that M_r and M_l have the following forms:

$$M_r = \begin{bmatrix} 0 & 0 \\ B_r & C_r \end{bmatrix}, \qquad M_l = \begin{bmatrix} 0 & B_l \\ 0 & C_l \end{bmatrix}.$$

Then the conditions (2) imply $A_{22}B_r = A_{21}$, $A_{22}C_r = A_{22}$, $B_lA_{22} = A_{12}$, and $C_lA_{22} = A_{22}$. The first and third of these equations obviously give us the conditions of Generalization IV, and hence we are done. ■

The following proposition may be found in [2].

Proposition 4. *If A is a positive matrix, then $S_{\mathrm{III}}(A) = S_{\mathrm{IV}}(A)$.*

Proof. First note that A positive implies that $\ker A_{22} \subseteq \ker A_{12}$ and hence $(\ker A_{12})^{\perp} = \operatorname{range} A_{12}^{*} \subseteq \operatorname{range} A_{22}^{*} = (\ker A_{22})^{\perp}$. Since $A_{12}^{*} = A_{21}$ and $A_{22}^{*} = A_{22}$, we see that the hypotheses of Generalization IV are met. Now, since $S_{\mathrm{IV}}(A) = S_{\mathrm{II}}(A)$, we see that $\operatorname{range} S_{\mathrm{IV}}(A) \subseteq S$. Let $\begin{bmatrix} c \\ y \end{bmatrix}$ be arbitrary, and define $z = y + A_{22}^{\dagger}A_{21}c$. Then

$$\left(A\begin{bmatrix} c \\ y \end{bmatrix}, \begin{bmatrix} c \\ y \end{bmatrix}\right) = \left(S_{\mathrm{IV}}(A)\begin{bmatrix} c \\ 0 \end{bmatrix}, \begin{bmatrix} c \\ 0 \end{bmatrix}\right) + (A_{22}z, z)$$

$$= \left(S_{\mathrm{IV}}(A)\begin{bmatrix} c \\ y \end{bmatrix}, \begin{bmatrix} c \\ y \end{bmatrix}\right) + (A_{22}z, z).$$

This shows $S_{\mathrm{IV}}(A) \leqslant A$, and moreover, since for each c we can choose y so as to make $z = 0$, we must have $S_{\mathrm{IV}}(A) \geqslant 0$. Thus $S_{\mathrm{IV}}(A) \in \mathscr{M}(A, S)$. Now suppose $B \in \mathscr{M}(A, S)$. Then since $\operatorname{range} B \subseteq S$, and B is positive, B must have the form

$$B=\begin{bmatrix} B_{11} & 0 \\ 0 & 0 \end{bmatrix}.$$

Thus

$$\begin{aligned}\left(B\begin{bmatrix} c \\ y \end{bmatrix},\begin{bmatrix} c \\ y \end{bmatrix}\right)&=\left(B\begin{bmatrix} c \\ -A_{22}^{\dagger}A_{21}c \end{bmatrix},\begin{bmatrix} c \\ -A_{22}^{\dagger}A_{21}c \end{bmatrix}\right)\\ &\leqslant\left(A\begin{bmatrix} c \\ -A_{22}^{\dagger}A_{21}c \end{bmatrix},\begin{bmatrix} c \\ -A_{22}^{\dagger}A_{21}c \end{bmatrix}\right)\\ &=\left(S_{IV}(A)\begin{bmatrix} c \\ y \end{bmatrix},\begin{bmatrix} c \\ y \end{bmatrix}\right).\end{aligned}$$

Therefore $B \leqslant S_{IV}(A)$ and $S_{IV}(A)$ is maximal in $M(A,S)$. ∎

PROPOSITION 5. *Generalizations* IV *and* V *are equivalent.*

Proof. First assume that the conditions of Generalization V are satisfied. Then since $A_{21}=XD$, $A_{22}=XY$, and $A_{12}=CY$, we see that $\ker X\subseteq\ker C$ implies $\ker A_{22}\subseteq\ker A_{12}$, and $\operatorname{range} D\subseteq\operatorname{range} Y$ implies $\operatorname{range} A_{21}\subseteq\operatorname{range} A_{22}$. Thus the conditions of Generalization IV are satisfied. We need only show now that $S_{IV}(A)=S_V(A)$. For this purpose choose $c\in S$, $y\in S^{\perp}$. Then there is a $x\in S^{\perp}$ and a $z\in S$ with

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} c \\ x \end{bmatrix}=\begin{bmatrix} z \\ 0 \end{bmatrix}.$$

Moreover

$$S_{IV}(A)\begin{bmatrix} c \\ y \end{bmatrix}=\begin{bmatrix} z \\ 0 \end{bmatrix}.$$

It now follows that $A_{21}c+A_{22}x=0$, or $X(Dc+Yx)=0$. Since $\ker X\subseteq\ker C$, we have $C(Dc+Yx)=0$ or $A_{12}x=-CDc$. Now

$$\begin{aligned}S_{IV}(A)\begin{bmatrix} c \\ y \end{bmatrix}=\begin{bmatrix} z \\ 0 \end{bmatrix}&=\begin{bmatrix} A_{11}c+A_{12}y \\ 0 \end{bmatrix}\\ &=\begin{bmatrix} (A_{11}-CD)c \\ 0 \end{bmatrix}=S_V(A)\begin{bmatrix} c \\ y \end{bmatrix}.\end{aligned}$$

Conversely, suppose the conditions of Generalization IV are satisfied. Let X be the projection onto the range of A_{22}, and let Y be A_{22}. Since $\operatorname{range} A_{21} \subseteq \operatorname{range} A_{22}$, if we let D be A_{21}, we have $A_{22} = XY$ and $A_{21} = XD$ with $\operatorname{range} D \subseteq \operatorname{range} Y$. Now by an elementary argument of linear algebra $\ker A_{22} \subseteq \ker A_{12}$ implies the existence of an operator C such that $A_{12} = CA_{22}$ or $A_{12} = CY$ with $\ker Y \subseteq \ker C$. (For the infinite dimensional version of this see [11].) Since we now have a proper decomposition, the first half of the proof shows

$$S_{\mathrm{IV}}(A) = \begin{bmatrix} A_{11} - CD & 0 \\ 0 & 0 \end{bmatrix}. \qquad \blacksquare$$

In summary, Generalizations II, IV, V, and VI are equivalent. If A is positive, then $S_{\mathrm{III}}(A) = S_i(A)$ for $i = \mathrm{I}, \mathrm{II}, \dots, \mathrm{VI}$. In any case, if S_i exists for $i = \mathrm{II}, \mathrm{III}, \dots,$ or VI, then $S_i(A) = S_{\mathrm{I}}(A)$.

REFERENCES

1 A. Albert, Conditions for positive and nonnegative definiteness in terms of pseudoinverses, *SIAM J. Appl. Math.* 17:434–440 (1969).
2 W. N. Anderson, Shorted operators, *SIAM J. Appl. Math.* 20:520–525.
3 W. N. Anderson, T. D. Morley, and G. E. Trapp, Cascade addition and subtraction of matrices, *SIAM J. Algebraic Discrete Methods*, to appear.
4 W. N. Anderson and G. E. Trapp, Shorted operators II, *SIAM J. Appl. Math.* 28:60–71 (1975).
5 T. Ando, Generalized Schur complements, *Linear Algebra Appl.* 27:173–186 (1979).
6 C. A. Butler and T. D. Morley, A note on the shorted operator, *SIAM J. Mat. Anal. Appl.* 9:147–155 (1988).
7 D. Carlson, What are Schur complements, anyway? *Linear Algebra Appl.* 74:257–275 (1986).
8 D. Carlson, Generalized inverse invariance, partial orders, and rank minimization problems for matrices, in *Current Trends in Matrix Theory*, ed. F. Uhlig and R. Grone, North Holland 1987.
9 D. Carlson, E. Haynsworth, and T. Markham, A generalization of the Schur complement by means of the Moore-Penrose inverse, *SIAM J. Appl. Math.* 26:169–179 (1974).
10 R. W. Cottle, Manifestations of the Schur complement, *Linear Algebra Appl.* 8:169–179 (1974).
11 R. Douglas, On majorization, factorization and range inclusion of operators on Hilbert space, *Proc. Amer. Math. Soc.* 17:413–416 (1966).
12 R. J. Duffin and T. D. Morley, Almost definite operators and electromechanical systems, *SIAM J. Appl. Math.* 35:21–30 (1978).

13 P. A. Filmore and J. P. Williams, On operator ranges, *Adv. in Math.* 7:254–281 (1971).
14 W. L. Green and T. D. Morley, Operator means, fixed points, and the norm convergence of monotone approximants, *Math. Scand.*, 60:202–218 (1987).
15 M. G. Krein, The theory of self adjoint extensions of semibounded Hermitian transformations and its applications I, II, *Mat. Sb.* 20:431–495 (1947); 21:205–224 (1947).
16 S. K. Mitra, Simultaneous diagonalization of rectangular matrices, *Linear Algebra Appl.* 47:139–150 (1982).
17 S. K. Mitra and M. L. Puri, Shorted matrices: An extended concept and some applications, *Linear Algebra Appl.* 42:57–79 (1982).
18 T. D. Morley, Several applications of the shorted operator, in *Constructive Approaches to Mathematical Models* (C. V. Coffman and G. Fix, Eds.), Academic.
19 D. V. Ouellette, Schur complements and statistics, *Linear Algebra Appl.* 36:187 (1981).
20 C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*, Wiley, New York, 1971.
21 C. R. Rao, S. K. Mitra, and P. Bhimasankaram, Determination of a matrix by its subclasses of generalized inverses, *Sankhyā Ser. A* 34:5–8 (1973).
22 I. Schur, Über Potenzreihen die im Innern des Einheitskreises beschrankt sind, *J. Reine Angew. Math.* 147:205–232 (1917).
23 J. J. Sylvester, On the relation between the minor determinants of linearly equivalent quadratic functions, *Philos. Magazine* 14:295–305 (1851).
24 G. E. Trapp, Hermitian semidefinite matrix means and related matrix inequalities —an introduction, *Linear and Multilinear Algebra* 16:113–123 (1984).

Received August 1986; final manuscript accepted 21 October 1987